For office use only	Team Control Number	For office use only
T1 ________	**2021050**	F1 ________
T2 ________		F2 ________
T3 ________		F3 ________
T4 ________		F4 ________

2021

The International Mathematical Modeling Challenge (IM^{2}C) Summary Sheet

(Your team's summary should be included as the first page of your electronic submission.)

Problem:

Today sport with its rich history, numerous kinds, competitions and athletes is one of the most exciting and topical aspects of our lives. Athletes train, develop their skills, take part in competitions to prove their mastery. Fans and spectators monitor the news, the tournaments, the stars in the industry relentlessly. But there is one thing that everyone is concerned about - who is the best athlete? This is the question that the GOAT award - Greatest of All Time - gives answer on.

How GOAT of any sport can be defined as objectively as possible? What algorithms and formulas need to be used to determine the best player in the competition?

Methods:

In our algorithm, we use paired comparison models (when participants compete in a series of personal competitions) and extend these models to take into account the time-changing strengths of competitors. That's why we consider the Bradley-Terry model. We also use barycentric rational interpolation that provides more accurate match, and our own methods that are characterized by simple mathematical considerations. All the data we receive are included in our final rating formula, which is based on the known formula for calculating the rating with the Elo method.

Results:

Key attributes have been identified for the mastery of athletes. On the basis of them a mathematical model was constructed for determining the rank of an athlete - an indicator that defines skills, abilities, talent, experience etc., for individual sports with competition in pairs. On the basis of the results of this ranking we nominate the person who deserves GOAT. The model was tested at the Summer Olympics in 2012, and specifically in badminton. Using the Python programming language and the tournament data, we get a ranking for each player and a final scoreboard. The model was expanded and adapted for all individual and team sports. Both strengths and weaknesses of the model were identified, as well as ways to address them.

Conclusion:

A mathematical model was developed and adapted for all sports to define an athlete worthy of the GOAT title. Its efficiency and effectiveness were tested by examples, and the results were consistent with both expert judgement and existing ratings. Suggestions were made on how to improve and develop the model, which in the future would allow a lot of standard and key factors to be taken into account.

Table of contents

1. Letter

Dear director of the Top Sport,

In an attempt to find a solution to the question you are interested in, we have developed a mathematical model with which you can determine the grandest player of all time.

Now let us move to the model itself. We have divided all sports into three categories: individual sports «one-on-one», individual round-robin sports and group sports. For each category, a formula was derived for calculating the rating of players, according to which GOAT is determined.

In individual one-on-one sports, the main indicator of a player's skill is the gap in points from the opponent in each match and, accordingly, victory/defeat. Also, the final rating of the player takes into account the strength of his opponent.

Similar to how we find the best athletes of individual sports «one-on-one», in individual round-robin sports, we compare all competitors in pairs to determine the final ranking,but instead of the gap on points, the key parameter is the gap on the corresponding sports parameter (for example, in running, this parameter is time).

The rating of players in group sports is already based on two components. Firstly, it is the contribution that the player made to the results of his team's match. Secondly, the rating of the team itself, which we found similar to the rating of players of individual sports «one-on-one».

We applied our model to two sports. For consideration, we have taken 4 major tennis competitions in 2018 (Wimbledon Championships, Australian Open, US Open, French Open). As a result, we found out that the leader of these tournaments was Simona Halep. We also reviewed the 2012 Olympic Badminton Games. Dan Lin became the leader there.

We are glad that you have contacted us with your problem.

Sincerely, team 2021050

2. Introduction

There are many vital aspects of our lives. Today sport with its rich history, numerous kinds, competitions and athletes is one of the most exciting and topical among them. For someone it's just a hobby, some people consider sport as a lifestyle, some just watch it from a television screen, and some people can't live without it. Athletes train, develop their skills, take part in competitions to prove their mastery. Fans and spectators monitor the news, the tournaments, the stars in the industry relentlessly. But there is one thing that everyone is concerned about - who is the best athlete? This is the question that the GOAT award - Greatest of All Time - gives answer on. This title was first pronounced by the now world-famous boxer Muhammad Ali. After one fight in 1964 a 22-year-old athlete first identified himself as the greatest athlete of all time. Since then every great athlete, without exception, has been dreaming of this unquestionably crucial title. But does everyone deserve it?

How GOAT of any sport can be defined as objectively as possible? What algorithms and formulas need to be used to determine the best player in the competition?

These are the questions that have become the basis of our research. We have been looking for and giving answers on them in order to develop a model that allows us to define an athlete worthy of the title of Greatest.

3. Task 1
3.1. Question a

3.1.1.
As we constructed the mathematical model, we identified the main indicators of the player's skill - his **point difference** (difference between his and opponent's points) in each set and, accordingly, his **win/loss**.
At first we wanted to use The Paired Comparison Method in the first task, but this was not the best solution, because if we recreate the matrix of these same comparisons, it would be incomplete. It is true that we cannot compare players who had not played together in the tournament.
For illustrative purposes, let's take an 8-person tournament with 4 quarter-finals (1-2,3-4,5-6,7-8), 2 semi-finals (1-3,5-7) and 1 final (1-5). That way the tournament system:

Player	1	2	3	4	5	6	7	8
1	-	+	+	-	+	-	-	-
2	+	-	-	-	-	-	-	-
3	+	-	-	+	-	-	-	-
4	-	-	+	-	-	-	-	-
5	+	-	-	-	-	+	+	-
6	-	-	-	-	+	-	-	-
7	-	-	-	-	+	-	-	+
8	-	-	-	-	-	-	+	-

Where the (+) means that the players had played against each other, and the (-) means that they had not. You can see that we can't compare in pairs a lot of players in this tournament.

3.1.2.
So we went the other way and came up with the following formula:

$$Pr_i = \sum_{t=1}^{k} Pr_t\left(1 + \frac{D}{Sum}\right),$$

where Pr_i - prestige/rank of the player i;
k - number of players that had lost to the player i. (Accordingly, Pr_k- rank of the player k);
D - difference between player's i and player's k points, i.e. how far the winner of the set has passed the opponent by points;
Sum - total points scored by both players during the set.

3.1.3. The structure of the model:
When one player defeats the other, his prestige is enhanced on the prestige of the other player considering their point difference. Thus, the rating of each player is composed of all ratings of the athletes that had lost to him.
In fact, using this model, we're looking at a tree where each row is the participants who have reached a certain stage, so at the very bottom - the participants who have lost immediately; and the highest vertex is the winner. Gradually moving from the bottom up, we find the ratings of all players and by them determine the best.

3.1.4. Notes:
- The formula is recursive, so the original $Pr_0 = const$ for those who lost first is needed to be set. We chose it so that the exponent part of the Pr_i for the next players would change only slightly. This way, the ranks could be fairly evaluated and compared.
- It is easy to notice that D may be negative or zero. However, it is obvious that $(1 + \frac{D}{S})$is positive in any case, therefore the prestige of the winner necessarily increases.

3.2. Question b

3.2.1.

In order to determine the best female tennis player of 2018, using the built model we examined 4 tournaments of Grand Slam and for each of them found ranks of the participants. After that, the gained ranks of each player in all tournaments were summed. Then on the created ranking table we determined the best athlete.

In this case, $Pr_0 = 0,1$.

Ranking tables:

US Open:

L Tsurenko		0,1151515			
K Kanepi	0,1				
M Vondrousova	0,1				
C Suares Navarro		0,1263158			
E Svitolina	0,1				
S Stephans		0,1333333			
M Keys			0,3095568		
E Mertens	0,1				
S Williams				0,788358	
A Barty	0,1				
A Sevastova			0,2907308		
N Osaka					1,7649926
M Sharapova	0,1				
A Sabalenka	0,1				
D Cibulkova	0,1				
Ka Pliskova		0,12			

Wimbledon Championships:

S-w Hsieh	0,1				
A Sasnovich	0,1				
A Van Uytvanck	0,1				
B Bencic	0,1				
D Vekic	0,1				
C Giorgi		0,1263158			
E Makarova	0,1				
Ka Pliskova	0,1				
J Ostapenko			0,303687		
J Gorges			0,2762414		
D Kasatkina		0,12			
S Williams				0,65366	
K Bertens		0,1181818			
D Cibulkova		0,1411765			
E Rodina	0,1				
A Kerber					1,5432162

Australian Open:

D Alleetova	0,1				
A Kontaveit	0,1				
N Osaka	0,1				
B Strycova	0,1				
M Rubarikova	0,1				
E Mertens			0,3745		
M Keys		0,1411765			
A Kerber			0,3392157		
C Garcia	0,1				
E Svitolina		0,175			
P Martic	0,1				
C Suares Navarro		0,1058824			
S Halep				0,67888	
C Wozniacki					1,4376469
S-w Hsieh	0,1				
Ka Pliskova		0,12			

French Open:

M Buzarnescu	0,1				
B Strycova	0,1				
M Sharapova		0,1			
A Kontaveit	0,1				
M Keys			0,2839682		
D Kasatkina		0,1181818			
E Mertens	0,1				
S Williams	0				
S Stephans				0,689463	
A Kerber		0,1411765			
C Garcia	0,1				
C Wozniacki	0,1				
G Muguruza			0,16		
S Halep					1,3508286
L Tsurenko	0				
Y Putintseva		0,1263158			

3.2.2. Final ranking table:

S Halep	2,029708105
A Kerber	2,023608397
N Osaka	1,864992568
C Wozniacki	1,537646931
S Williams	1,442018134
S Stephans	0,822796511
S Williams	0,788357914
M Keys	0,73470149
E Mertens	0,5745
Ka Pliskova	0,34
J Ostapenko	0,303687025
A Sevastova	0,290730838
J Gorges	0,276241406
E Svitolina	0,275
D Cibulkova	0,241176471
D Kasatkina	0,238181818
C Suares Navarro	0,232198142
S-w Hsieh	0,2
M Sharapova	0,2
A Kontaveit	0,2
B Strycova	0,2
C Garcia	0,2
G Muguruza	0,16
C Giorgi	0,126315789
Y Putintseva	0,126315789
K Bertens	0,118181818
L Tsurenko	0,115151515
A Sasnovich	0,1
A Van Uytvanck	0,1
B Bencic	0,1
D Vekic	0,1
E Makarova	0,1
E Rodina	0,1
K Kanepi	0,1
M Vondrousova	0,1
A Barty	0,1
A Sabalenka	0,1
D Alleetova	0,1
M Rubarikova	0,1
P Martic	0,1
M Buzarnescu	0,1

3.2.3 Result:

The best female tennis player of 2018 - Simona Halep

Note: You can also use different coefficients that take into account the importance of tournaments; we considered them to be equal.

4. Task 2

4.1. Question a

4.1.1.

4.1.1.1.

We create the model based on the Elo rating system and the Barycentric Rational Interpolation. That way, our model is as follows:

$$Pr_i' = Pr_i + U(S_k - E_k)\ (1)$$
$$(+\ w = const\ in\ case\ of\ victory),$$

where Pr_i - rank of the player i before game/fight/competition, Pr_i' - after;

$U = const$ - the coefficient that we choose so that the exponent part of the Pr_i would change only slightly;

S_k - **actual** relative scoring in the game with player k. (Relative scoring - $\frac{D}{Sum}$(See details in task 1, question a));

E_k - **suggested** relative scoring in the game with player k.

Note: in case of a win we add $w = const$ so that if the player does not meet expectations, that is $S_k - E_k < 0$, the player's rank will increase anyway. We chose w so that the exponent part of the Pr_i would change only slightly.

4.1.1.2.

Now let's focus on **how we get** E_k.

We find it from the probability that in this competition participant i will defeat player k. So let's determine her first.

$$p_{ik} = \frac{a_i}{a_i + a_k},\quad (2)$$

where p_{ik} - required probability;

a_i, a_k - the strength of players i and k respectively and calculated at time t by means of a Barycentric Rational Interpolation as follows:

$$a_i(t) = \frac{\sum_{l=1}^{n_i} \omega_{il} a_{il}/(t-t_{il})}{\sum_{l=1}^{n_i} \omega_{il}/(t-t_{il})}, \quad (3)$$

where a_{il} -the strength of the player i at time t_{il};

ω_{il} - Weight for Rational Barycentric Interpolation. It was found that for $n_i \le 5$ weight of order zero should be taken: $\omega_l = (-1)^l$, and for $n_i \ge 5$ - weight of first order: $\omega_l = (-1)^l(\frac{1}{t_l - t_{l-1}} + \frac{1}{t_{l+1} - t_l})$ to best match the data;

n_i - the nodes throughout a player i's career, which are evenly distributed. However, we remove the node if it has no support in the form of played matches/competitions during this period of time. The final formula for n_i, determined experimentally:

$$n_i = 1 + y_i f_i + (y_i/M)(1 - f_i), \quad (4)$$

where $y_i + 1$ - the number of player i's played years;

$M = const = 4$ - the coefficient;

$$f_i = \left|\frac{1-exp(hg_i/y_i)}{1+exp(hg_i/y_i)}\right| \quad (5)$$

- a logical type function where g_i is the total number of games of an athlete. f_i is in range from 1 for athletes who never play, to 0 for athletes who play often, etc.

n_i progresses from 1 to 4 for athletes who played few matches, to annual nodes for those who played often.

The coefficient $h \simeq 0,05$ was introduced to minimize the information criterion (AIC).

In order to get E_k (the suggested relative scoring), we must divide the probability difference by their sum, but since the events are independent, then $p_{ik} + p_{ki} = 1$. So the formula $E_k = \frac{p_{ik} - p_{ki}}{p_{ik} + p_{ki}}$ can be introduced as $E_k = 2 * p_{ik} - 1$.

4.1.2.

After constructing the mathematical model, we implemented it by writing a program. We chose badminton as the object of our research.

4.1.2.1. What does the program do?

Since it would take a lot of time and data to define GOAT badminton, we will test our algorithm on the 2012 Summer Olympics. By using formulas (5) and (4) we will determine the number of nodes for each player depending on how many games and years they have played.

Then, using formula (3), we recursively calculate the player's strength just before the Olympics. After, using formula (2), we can calculate the probability of one player defeating another in a match, multiplying this by 2 and then reducing it by 1. The resulting value will be the suggested relative scoring in the game with the opponent.

We further assume that the initial rank of each player is 100 (similar to the fact that at the beginning of each player's career he has some initial rank). After that, we'll keep the difference between the points in the sets and who won the game in the individual list. Then, we calculate a new rank of the player after winning the match according to formula (1), taking the coefficient U=50 to maintain the ranking order.

After performing these actions for each game we will get a ranking of all players of the tournament, so we won't only determine MVP Olympics, but we will also find out what places the other participants were in and with what ranks.

Can GOAT be defined by this algorithm?

Yes. At the beginning of a career, each player will have some fixed rating. Immediately before each match (tournament), we calculate the suggested relative scoring. After the game itself, we calculate his new rank by a formula. This is how we do it for all players in this sport and for their entire career. The one with the highest rank will be the GOAT of badminton.

Code fragment:

```
File  Edit  Format  Run  Options  Window  Help
import math as m
file=open("Data.txt")
text=open("Results.txt")
str=file.read().split()
result=text.read().split()
res=[]
g=[]
d=[]
power=[]
for i in range(16):
    res.append(int(result[i]))
for i in range(32):
 if i%2==0:
    d.append(int(str[i]))
 else:
    g.append(int(str[i]))
for i in range(16):
 f=abs((1-m.exp(0.05*g[i]/d[i]))/(1+m.exp(0.05*g[i]/d[i])))
 n=round(1+d[i]*f+(d[i]/4)*(1-f))
 a=[0,1,2]
 w=[0,-1]
 for k in range(3,n+1):
   if n<=5:
       w.append((-1)**k)
   else:
       w.append(((-1)**k)*2)
   summa1=0
   summa2=0
   for j in range(1,k):
       summa1=a[j]*w[j]/(k-j)+summa1
       summa2=w[j]/(k-j)+summa2
   a.append(round(summa1/summa2,3))
 for u in range(len(a)):
   a[u]=int(a[u]*1000)
 power.append(a[len(a)-1])
rating=[100,100,100,100,100,100,100,100,100,100,100,100,100,100,100,100]
for i in range(0,16,2):
 p=power[i]/(power[i]+power[i+1])
 rating[i]=round(rating[i]+50*(res[i//2]/42-(2*p-1)))
print(rating)
win=[2,4,6,10]
lose=[14,12,0,8]
```

Raw data:

Number	Name	Years	Games	1/8	1/4	1/2	1
1	Kashyap PARUPALLI	6	126	13			
2	Niluka KARUNARATNE	7	102				
3	Dan LIN	7	156	21	9	20	7
4	Taufik HIDAYAT	8	168				
5	Hyun Il LEE	9	161	12	11		
6	Jan Ostergaard JORGENSEN	7	142				
7	Chong Wei LEE	9	158	22	12	15	
8	Simon SANTOSO	6	109				
9	Peter GADE	11	180	17			
10	Wan Ho SON	6	138				
11	Long CHEN	6	126	8	13		9*
12	Wing Ki WONG	4	90				
13	Jin CHEN	8	112	19			
14	Marc ZWIEBLER	10	172				
15	Sho SASAKI	8	148	13			
16	Kevin CORDON	9	105				

4.1.2.2.

As a result, we receive each athlete's rank:

```
Kashyap PARUPALLI 115
Niluka KARUNARATNE 100
Dan LIN 169
Taufik HIDAYAT 100
Hyun Il LEE 128
Jan Ostergaard JORGENSEN 100
Chong Wei LEE 160
Simon SANTOSO 100
Peter GADE 121
Wan Ho SON 100
Long CHEN 133
Wing Ki WONG 100
Jin CHEN 122
Marc ZWIEBLER 100
Sho SASAKI 116
Kevin CORDON 100
```

So we see that Dan Lin has the highest rank. According to the results of the Olympics, he is the best.

As an example, we took one Olympics, but our model also works on long timescales, so we can look at the careers of athletes in full and calculate their final ranks. So we can define the best player of the 21st century, the 20th century, we can even compare athletes from different centuries, which doesn't always make sense, because the skills, the strategies of the games and so on change and are improved over time.

4.2. Question b

4.2.1.

Let's talk about how to adapt our model to all individual sports. For species where the competition is one-on-one, our model is adapted, the relative scoring (S_k, E_k) will be the **scoring of the respective sport parameter** and the formula remains the same.

Now let's look at round-robin species like running, where we don't have one-on-one games, and everyone competes against everyone.

In the event of a round-robin competition, we can consider all **pairs** of participants, reducing this case to our built model.

So the relative scoring (S_k, E_k) will be the scoring of the respective sport parameter too (for example, in running time is this parameter).

4.2.2.

Thus the final formula looks like this:

$$Pr_i' = Pr_i + \frac{\sum_{t=1}^{k} U(S_t - E_t)}{k},$$

where k - participants that compete against player i and all parameters are set in a similar way to the built model, coefficient U is chosen so that the exponent part of the Pr_i would change only slightly. Thus, it is easy to get ranks of athletes both for certain competitions and for the whole career, and by them to determine the best, worthy of the GOAT title .

5. Task 3

5.1.

Let's adapt our model for team sports. With its help, we can find the team's rank in the same way as the player's rank.

It remains to determine the player's rank based on the team's rank. To do this, we will introduce another parameter - The player's effectiveness in some particular game - η. Then the rank of the athlete will be calculated as follows:

$$Pr_i' = Pr_i + \eta_i \times \Delta R ,$$

where ΔR - changing of the rank of a player's i team for a given game (it can be either positive or negative);

Pr_i' - rank of the player i after this game;

Pr_i - rank of the player i before this game;

η_i - effectiveness of the player i.

5.2.

Now let's figure out how to calculate this effectiveness:

It is important to notice that when a team wins, the better the player plays (that is, the more points he has), the greater their efficiency is. However, if a team loses, the rank change ΔR is negative, consequently, the fewer points a player gets per game, the more he "helps" to lose, and the greater the coefficient ΔR must be to recalculate his rank.

So, the effectiveness of the player for the game won by his team, we consider as:

$$\eta_i = \frac{q_{i+1}}{\sum_{t=1}^{l} q_t + n},$$

And the effectiveness of the player from the losing team as:

$$\eta_i = 1 - \frac{q_{i+1}}{\sum_{t=1}^{l} t + n},$$

where q_l - the number of points scored by player l per game;

l - players of the same team;

n- number of players per team.

Thus, the effectiveness of each player is a positive number less than one.

5.3.

The effectiveness in fact depends on how many points the team and a particular player scored. Let's see how to count these points. For each sport, they need to be defined differently, but the concept remains the same: the characteristic of these points is the worth of the athlete per game, in most team sports it is measured by the number of interactions with other players. For example, for volleyball, these points will be awarded according to the following rules:

- if a successful transfer of the ball is made between two players of the same team, each is awarded 0.5 points;
- if an unsuccessful transfer is made between two players of the same team, no points are added;
- if the ball is passed between players of different teams and the ball does not fall, then each player gets 0.5 points;
- if the ball is passed between players of different teams and the ball falls, then the player who wins this transfer is awarded 1 point, and his opponent - 0.

So the total number of points of a team will be the sum of the points of all its players.

5.4.

When a player changes teams, the player's rank is saved.

The rule "the rank of a team is equal to the sum of the ranks of its players" is preserved, This means that when a player leaves a certain team, his rank is subtracted from the overall rank of this team, and added to the rank of the team in which he came.

Similarly, we can calculate such rank both for a certain game and for the whole career of the players, and then compare the athletes by these ranks and determine who is worthy of the GOAT title.

6. Results:

6.1.

The main differences and advantages of our model:

- No subjective factors affecting the rating (such factors are present in figure skating, etc.);
- Our model takes into account the strength of the players competing with each other. This allows us to fairly distribute the points among the players. Since for an experienced and strong athlete A it won't be difficult to defeat a beginner B. For this victory, the rank of athlete A can't increase much. And the rank of the athlete B won't fall much;
- Accounting of all games in which the player participated (unlike, for example, the scoring system of "ATP Rankings", which takes into account the points scored by a player only for the top 18 tournaments in a season). This allows us to evaluate not only the successful games of the athlete, but also to take into account all the mistakes which he made;
- Regardless of which score the game ended with, the player rating changes (in contrast to the ranking of the ABA (Amateur Badminton Association)), where, if the rank of the winning player is greater than the rank of the losing player by more than 100 points, then points are not assigned to the stronger player for the victory, and are not deducted from the weaker player) This is the most fair to the players. Since if a weaker athlete decides to take a risk and try himself at the leading competitions, then he must pay for his choice. At the same time, a strong player should not suffer because of this choice.
- Our model makes it possible to evaluate and compare athlete in different periods of his career; athletes who have never played with each other and even those who lived in different centuries (however, it is also necessary to understand that this doesn't always make sense, as the sport and its elements, strategies, and skills change and develop).
- Our model makes it possible not only to compare different players, but also teams for group sports.

6.2.

Disadvantages of the model:

The main drawback of our model is that for each sport you need to select your own parameters U, Pr_0, etc. Moreover, you can only select these parameters experimentally or through complex analysis.

7. Ways to improve:

It is easy to notice that in addition to the final score in various sports, some game moments also affect the athlete's rank. For example, in chess, the fewer moves were made before checkmate, the better the winning athlete played, the worse the losing athlete played, and vice versa. In basketball, there are many different offenses, such as hand checking or double dribble. In this case, the player's personal rank should be reduced.

Therefore, it is necessary to develop a system of allowances and deductions for various expressions of the player. It will be based on the principles of normal distribution, since the probability of getting the minimum or maximum for some criteria is very low. For example, for the described example with chess. Or on the principles of asymmetrical distribution for quantities like the number of fouls in basketball. For no fouls, the player won’t receive anything, and for a large number of fouls, he will receive a strong deduction.

So, in individual disciplines, we will be able to introduce a system of rewards for some record-breaking performances and encourage athletes to play more honestly and carefully. In the team disciplines, we will be able to separate the player’s rank from the team’s rank even more, since outstanding players can play in weak teams and vice versa.

8. Conclusion

In the course of the work, we have determined the main parameters that characterize the skill of athletes. Taking into consideration these parameters we have developed and constructed a mathematical model which can give the athlete a rank. That rank defines skills, abilities, talent, experience, etc and can be adapted for each type of sport. As a result, we can choose sportsmen who deserve to be named GOAT. The efficiency and productiveness of the model was tested using a Python program with examples. The results were consistent with both the subjective assessments of experts and the existing ranks.

We have found strengths and weaknesses of the model, and also ways to improve it. Ways to improve and develop the model were suggested, which in the future will allow us to take into account as many factors as possible.

9. Literature

[1] A dynamic paired comparisons model: Who is the Greatest Tennis Player?, Rose D.Bakerlan, G. McHale.

[2] Deterministic Evolution of Strength in Multiple Comparisons Models: Who is the Greatest Golfer?, Rose D.Baker, Ian G. McHale.

[3] Who was the greatest of all-time? A historical analysis by a complex network of professional boxing, Adam G. Tennant, Chase M. L. Smith, Jotam E. Chen C.

[4] Who Is the Best Player Ever? A Complex Network Analysis of the History of Professional Tennis, Filippo Radicchi.

[5] G.O.A.T. Metrics. Who’s the Greatest Of All the Time?

[6] Greatest Tennis Players of all time "GOATs"

[7] Who is the GOAT: Jordan, Bryant, or King James? An Inference Based on Data Crunching of the Surface Web, Ahmed Al-Imam, J Athl Enhanc 2017

[8], [9] Creating An NBA G.O.A.T. Pyramid, JxmyHighroller

[10] Elo rating system, Wikipedia The Free Encyclopedia

[11] PageRank, Wikipedia The Free Encyclopedia
[12] Information criterion, Wikipedia The Free Encyclopedia